

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 16, 2003, 15:50:54 ; Search time 2343.67 Seconds  
(without alignments)  
16465.781 Million cell updates/sec

Title: US-09-497-967-1

Perfect score: 1326

Sequence: 1 atgaataataattttatt.....ttattttttctttttttg 1326

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_on:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_in:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1326	100.0	2486	3	AF140273	AF140273 Ichthyoph
2	1170.4	88.3	1249	3	ICVIMANT	M92907 Ichthyoph
3	647.2	48.8	1520	3	AF405431	AF405431 Ichthyoph
4	252.6	19.0	3026	3	AF324424	AF324424 Ichthyoph
C 5	94	7.1	131274	2	AC096869	AC096869 Rattus no
C 6	94	7.1	132449	9	AL365272	AL365272 Human DNA
C 7	94	7.1	176822	2	AC099174	AC099174 Rattus no
C 8	93	7.0	179553	2	AC024253	AC024253 Homo sapi
C 9	89.8	6.8	170985	2	AC096032	AC096032 Rattus no
C 10	88.8	6.7	186935	2	AC022322	AC022322 Homo sapi
C 11	88.4	6.7	189461	2	AC119627	AC119627 Rattus no
C 12	88.4	6.7	191841	2	AC121374	AC121374 Rattus no
C 13	87.6	6.6	204259	2	AC110817	AC110817 Mus muscu
C 14	87.6	6.6	264522	2	AC090437	AC090437 Mus muscu
C 15	86.2	6.5	84472	2	AC096684	AC096684 Takifugu
C 16	83.8	6.3	1611	10	MMU70651	U70651 Mus musculu
C 17	82.6	6.2	172307	2	AC044842	AC044842 Homo sapi
C 18	82.6	6.2	201470	2	AC113649	AC113649 Rattus no
C 19	82.4	6.2	149228	2	AC113774	AC113774 Rattus no
C 20	82.4	6.2	261604	2	AC119819	AC119819 Mus muscu
C 21	82.2	6.2	165318	2	AC126747	AC126747 Rattus no
C 22	81.8	6.2	175699	2	AC129136	AC129136 Rattus no
C 23	81.6	6.2	131346	2	AC119558	AC119558 Rattus no
C 24	80	6.0	88203	5	AC097628	AC097628 Takifugu
C 25	79.4	6.0	176774	2	AC109366	AC109366 Rattus no
C 26	79	5.9	209973	2	AC073755	AC073755 Mus muscu
C 27	78.2	5.9	71553	2	AC096997	AC096997 Takifugu
C 28	78.2	5.9	175748	2	AC120669	AC120669 Rattus no
C 29	78.2	5.9	180903	2	AC125859	AC125859 Rattus no
C 30	77.4	5.8	1612	10	MMU70652	U70652 Mus musculu
C 31	77.4	5.8	196650	2	AC111217	AC111217 Rattus no
C 32	77.2	5.8	85786	9	AL162582	AL162582 Human DNA
C 33	77	5.8	186380	9	AC100822	AC100822 Homo sapi
C 34	76.6	5.8	173540	2	AC119697	AC119697 Rattus no
C 35	76.6	5.8	180668	2	AC020857	AC020857 Mus muscu
C 36	76.2	5.7	183413	2	AC131200	AC131200 Rattus no
C 37	76	5.7	168757	2	AC127508	AC127508 Homo sapi
C 38	75.6	5.7	87201	9	AC007623	AC007623 Homo sapi
C 39	75.4	5.7	172853	9	AC084361	AC084361 Homo sapi
C 40	75.4	5.7	185994	2	AC002042	AC002042 Homo sapi
C 41	75.2	5.7	186558	2	AC079031	AC079031 Homo sapi
C 42	75	5.7	216977	2	AC105611	AC105611 Rattus no
C 43	75	5.7	219761	2	AC105642	AC105642 Rattus no
C 44	74.8	5.6	325808	2	AC121787	AC121787 Mus muscu
C 45	74.2	5.6	108476	2	AC094507	AC094507 Rattus no

# ALIGNMENTS

RESULT 1  
AF140273  
LOCUS  
DEFINITION Ichthyophthirius multifiliis immobilization antigen precursor  
ACCESSION AF140273  
VERSION AF140273.1 GI:4868370  
KEYWORDS  
SOURCE Ichthyophthirius multifiliis  
ORGANISM Ichthyophthirius multifiliis  
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;  
Hymenostomatida; Ophryoglenina; Ichthyophthirius.  
REFERENCE  
AUTHORS 1 (bases 1 to 2486)  
Clark,T.G., Lin,T.L., Jackwood,D.A., Sherrill,J., Lin,Y. and  
Dickerson,H.W.

The gene for an abundant parasite coat protein predicts tandemly  
 repetitive metal binding domains  
 Gene 229 (1-2), 91-100 (1999)  
 MEDLINE  
 99196987  
 PubMed  
 10095108  
 2 (bases 1 to 2486)  
 Gaertig,J., Gao,Y., Tishgarten,T., Clark,T.G. and Dickerson,H.W.  
 Surface display of a parasite antigen in the ciliate Tetrahymena  
 thermophila  
 Nat. Biotechnol. (1999) In press  
 JOURNAL  
 3 (bases 1 to 2486)  
 Clark,T.G., Lin,T.-L., Jackwood,D.A. and Dickerson,H.W.  
 Direct Submission  
 Submitted (29-MAR-1999) Microbiology & Immunology, Cornell  
 University, College of Veterinary Medicine, Ithaca, NY 14853, USA  
 JOURNAL  
 Location/Qualifiers  
 1..2486  
 /organism="Ichthyophthirius multifiliis"  
 /serotype="A"  
 /isolate="G1"  
 /db\_xref="taxon:5932"  
 399..1845  
 /gene="iAG48"  
 399..1845  
 /gene="iAG48"  
 /product="immobilization antigen precursor"  
 433..1761  
 /gene="iAG48"  
 /note="surface protein; 48 kDa i-antigen"  
 /codon\_start=1  
 /transl\_table=6  
 /product="immobilization antigen precursor"  
 /protein\_id="RAD31283.1"  
 /db\_xref="GI:4868371"  
 /translation="MKYINLLIILISIFINELRAVPCPDGTQAGLTDVGAADLGTG  
 VNCRFNYNGAAGQVFAANNAARGICVPCOINRVGSVFNAGDAILATGTC  
 STOCPTGALDDGVDFDRSAQCKVKPNFYNGSPQGEAPGVQVFAAGAAAGV  
 AAVTSQVPCOLKNDSPATAGANLATQCSNOCPTGVLDDGTVLTFNFTSLCVK  
 CRPNFYNGSPQGEAPGVQVFAAGAAAGVAAVTSQVPCOINRVNDSPATAGANL  
 ATQCSQPTGTAIDGTVLTFNFTSLCVKCRPNFYNGSPQGEAPGVQVFAAGAAAGV  
 PAHAPGNTATQATQCLTTCBPAGTVLDDGTSNFAVASECTKCSAGFEASKTGFTAG  
 PDTCCTCKLTSAGTAKVYAEATQKVOCASTTFAKLISLISLIFISYLL"  
 433..492  
 /gene="iAG48"  
 493..1758  
 /gene="iAG48"  
 /product="immobilization antigen"  
 BASE COUNT 896 a 310 c 321 g 959 t  
 ORIGIN  
 Query Match 100.0%; Score 1326; DB 3; Length 2486;  
 Best Local Similarity 100.0%; Pred. No. 6,1e-241;  
 Matches 1326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 ATGAATATATATATTTTATTAATTTTCTTTTATTTATTTATTAATGAATAGAGCT 60  
 Db 433 ATGAATATATATTTTATTAATTTTCTTTTATTTATTTATTTATTAATGAATAGAGCT 492  
 61 GTTCCATGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120  
 Db 493 GTTCCATGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552  
 121 CTGGTACTTGTGTAATTCGACACCTAAATTTTACTATAATGGTGGTGGTGGTGGT 180  
 Db 553 CTGGTACTTGTGTAATTCGACACCTAAATTTTACTATAATGGTGGTGGTGGTGGT 612  
 181 GAAGCTAATGTAATTAACCTTCGAGAGAAATTAATGCTGCTAGAGGTATATGCTACCA 240  
 Db 613 GAAGCTAATGTAATTAACCTTCGAGAGAAATTAATGCTGCTAGAGGTATATGCTACCA 672  
 241 TGCCAAATAAAGAGTAGGCTGTTTACCAATGCAGGTGACTTAGCTACTTTAGCCACA 300  
 Db 673 TGCCAAATAAAGAGTAGGCTGTTTACCAATGCAGGTGACTTAGCTACTTTAGCCACA 732

QY 301 TAATGACGACTTAATGTCCTTACTGGCAGCTGCACCTTGATGATGAGTGACAGATGTTTT 360  
 Db 733 TAATGACGACTTAATGTCCTTACTGGCAGCTGCACCTTGATGATGAGTGACAGATGTTTT 792  
 QY 361 GATAGATCAGCGGATTAATGTTAAATGCAAACTAACTTTTACTATAATGTTGTTCT 420  
 Db 793 GATAGATCAGCGGATTAATGTTAAATGCAAACTAACTTTTACTATAATGTTGTTCT 852  
 QY 421 CCTTAAGCTGAAGCTCCTGGCGTTTAAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 Db 853 CCTTAAGCTGAAGCTCCTGGCGTTTAAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
 QY 481 GCTGCGCTTACTAGTTAATGTTACCTTGGCAACTAAACAAAAACGATTTCTCTGCCACT 540  
 Db 913 GCTGCGCTTACTAGTTAATGTTACCTTGGCAACTAAACAAAAACGATTTCTCTGCCACT 972  
 QY 541 GCAGGTGCTTAAGCTTAATTTAGCCACATAATGTAGCAATTAATGCTTACTTGGCAGTGA 600  
 Db 973 GCAGGTGCTTAAGCTTAATTTAGCCACATAATGTAGCAATTAATGCTTACTTGGCAGTGA 1032  
 QY 601 CTTGATGATGAGTGACACTTGTGTTTAAATACATCAGCCPACATTTATGTTAAATGACGA 660  
 Db 1033 CTTGATGATGAGTGACACTTGTGTTTAAATACATCAGCCPACATTTATGTTAAATGACGA 1092  
 QY 661 CCTAAGCTTTACTATAATGTTGCTTCTTAAAGTGAAGCTCCTGGCGTTTAAAGTTTT 720  
 Db 1093 CCTAAGCTTTACTATAATGTTGCTTCTTAAAGTGAAGCTCCTGGCGTTTAAAGTTTT 1152  
 QY 721 GCTGCTGTTGCTGCGCTGACAGTGTGCTGCGTTTACTAGTTAATGTTGCTTGGCAA 780  
 Db 1153 GCTGCTGTTGCTGCGCTGACAGTGTGCTGCGTTTACTAGTTAATGTTGCTTGGCAA 1212  
 QY 781 ATAAACAAAAAGATTTCTCTCCACTGCGAGTGCCTAAGCTAATTTAGCCACATAATGC 840  
 Db 1213 ATAAACAAAAAGATTTCTCTCCACTGCGAGTGCCTAAGCTAATTTAGCCACATAATGC 1272  
 QY 841 AGTACTTAATGTCCTCAACTGGCACTGCAATTCAGAGGAGTGACACTGTTGTTTAGTAAT 900  
 Db 1273 AGTACTTAATGTCCTCAACTGGCACTGCAATTCAGAGGAGTGACACTGTTGTTTAGTAAT 1332  
 QY 901 TCATCCACATAATGTTCTTAAATGCAATGCTAATTTTAAATGTTAATTTGGAAGCA 960  
 Db 1333 TCATCCACATAATGTTCTTAAATGCAATGCTAATTTTAAATGTTAATTTGGAAGCA 1392  
 QY 961 GGTAAAGTTAATGTTTAAAGTGCAGTAAGTAAACTTCCAGACATGCTCCAGCT 1020  
 Db 1393 GGTAAAGTTAATGTTTAAAGTGCAGTAAGTAAACTTCCAGACATGCTCCAGCT 1452  
 QY 1021 AATAGTCTTAAAGCCACATAATGTTTGAACACATGCTGCTGGTACAGTACTTGTAT 1080  
 Db 1453 AATAGTCTTAAAGCCACATAATGTTTGAACACATGCTGCTGGTACAGTACTTGTAT 1512  
 QY 1081 GATGGAACATCAACTAATTTTGTAGCTCCGCAACTGGAATGATGATGATGATGATGATGAT 1140  
 Db 1513 GATGGAACATCAACTAATTTTGTAGCTCCGCAACTGGAATGATGATGATGATGATGATGAT 1572  
 QY 1141 TTTTTCGATCAAAACAACTGGTTTACAGCAGGAGTACATGATGATGATGATGATGATGAT 1200  
 Db 1573 TTTTTCGATCAAAACAACTGGTTTACAGCAGGAGTACATGATGATGATGATGATGATGAT 1632  
 QY 1201 AAAAAATTAATCTTCTGGTGCACAGCTAAAGTATATGCTGAAGCTACTCAAAAAGTATAA 1260  
 Db 1633 AAAAAATTAATCTTCTGGTGCACAGCTAAAGTATATGCTGAAGCTACTCAAAAAGTATAA 1692  
 QY 1261 TGCGCTCCACTACTTTCGGTAAATTTTATCGATTTTCTTATTTATTTTCTTCTAT 1320  
 Db 1693 TGCGCTCCACTACTTTCGGTAAATTTTATCGATTTTCTTATTTATTTTCTTCTAT 1752  
 QY 1321 TTAATG 1326  
 Db 1753 TTAATG 1758

RESULT 2  
ICYMANT  
LOCUS  
DEFINITION  
ICthyophthirius multifiliis mRNA linear INV 18-SEP-1998  
mRNA, partial cds.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
M92907.1 GI:3628568  
ORGANISM  
Ichthyophthirius multifiliis.  
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;  
Hymenostomida; Ophryoglenina; Ichthyophthirius.  
REFERENCE  
AUTHORS  
TITLE  
Lin, T. L. and Dickerson, H. W.  
Purification and partial characterization of immobilization  
antigens from Ichthyophthirius multifiliis  
J. Protozool. 39 (4), 457-463 (1992)  
JOURNAL  
MEDLINE  
PUBMED  
93020590  
1383510  
REFERENCE  
AUTHORS  
TITLE  
Clark, T. G., McGraw, R. A. and Dickerson, H. W.  
Developmental expression of surface antigen genes in the parasitic  
ciliate Ichthyophthirius multifiliis  
Proc. Natl. Acad. Sci. U.S.A. 89 (14), 6363-6367 (1992)  
JOURNAL  
MEDLINE  
PUBMED  
92335298  
1631132  
REFERENCE  
AUTHORS  
TITLE  
Clark, T.  
Direct Submission  
Submitted (06-JUN-1992) Microbiology and Immunology, Cornell  
University, Ithaca, NY 14853, USA  
JOURNAL  
AUTHORS  
TITLE  
Clark, T.  
Direct Submission  
Submitted (18-SEP-1998) Microbiology and Immunology, Cornell  
University, Ithaca, NY 14853, USA  
REMARK  
COMMENT  
Sequence update by submitter  
On Sep 18, 1998 this sequence version replaced gi:159289.  
FEATURES  
source  
1. .1249  
Location/Qualifiers  
/organism="Ichthyophthirius multifiliis"  
/strain="Georgia"  
/db\_xref="taxon:5932"  
/dev\_stage="trophont"  
<1..1190  
/note="48 kDa"  
/citation=[2]  
/citation=[1]  
/codon\_start=3  
/evidence=experimental  
/transl\_table=6  
/product="immobilization antigen precursor"  
/protein\_id="AAC36158.1"  
/db\_xref="gi:3628569"  
/translation="AVPCPDGTQAGLDVGAADLGTCVNCRPNYNGGAAQGSAN  
GNQFPAANNAARGICVPCQINRVGSVTVNAGDLATLATQCSTQCTGTALDDGVTVDFD  
RSAQGVCKRPNFYNGSPQGEAPGVQVFAAGAAAGAAVAVTSQVPCOLNKNDSPA  
TAGAANLATQCSNQCTPTVLDGVTLVENTSATLCKRPNFYNGSPQGEAPGV  
QVFAAGAAAGAAVAVTSQVPCQINKNDSPATAGAOANLATQCSQCTPTGTALQDGV  
LVFNSNSTQCSQCTANYFENGLEAGSQCLKCPVSKTTPFAHPAGNATQATQCLTFC  
PAGVLDGDTSTNFASATETKCSAGFFASKTGTGFTAGDTCTECTKLTSGATAVN  
IHQ"  
sig\_peptide  
mat\_peptide  
6..1187  
/product="immobilization antigen"  
/evidence=experimental  
177..893  
/citation=[2]  
repeat\_region  
/rpt\_type=tandem  
/rpt\_unit=414..680  
BASE COUNT 348 a 243 c 257 g 401 t  
ORIGIN

Query Match 88.3%; Score 1170.4; DB 3; Length 1249;  
Best Local Similarity 99.9%; Pred. No. 1.8e-211;  
Matches 1171; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 56 GAGCTGTTCCATGTCCTGATGGTACTTAGACTCAAGCTGGATGACTGATGAGTGCTG 115  
Db 1 GAGCTGTTCCATGTCCTGATGGTACTTAGACTCAAGCTGGATGACTGATGAGTGCTG 60  
QY 116 CTGATCTTGGTACTTGTGTTAATTCGAGACCTAAATTTTACTATAATGCTGGTGGCTG 175  
Db 61 CTGATCTTGGTACTTGTGTTAATTCGAGACCTAAATTTTACTATAATGCTGGTGGCTG 120  
QY 176 AAGGAGAAGCTAATGGTAATTAACCTTTTCGAGCAAAATATGCTGCTAGAGTATATG 235  
Db 121 AAGGAGAAGCTAATGGTAATTAACCTTTTCGAGCAAAATATGCTGCTAGAGTATATG 180  
QY 236 TACCATGCCAAATAAAGACAGTAGGCTCTGTACCAATTCAGGTGACTTTAGCTACTT 295  
Db 181 TACCATGCCAAATAAAGACAGTAGGCTCTGTACCAATTCAGGTGACTTTAGCTACTT 240  
QY 296 CCACATAATGACGTACTTAATGTCCTACTGGCACTGCACCTTGATGATGGAGTGACAGAT 355  
Db 241 CCACATAATGACGTACTTAATGTCCTACTGGCACTGCACCTTGATGATGGAGTGACAGAT 300  
QY 356 TTTTGTGATAGATCAGCCGCATATGTTAAATGCAAACTAACTTTTACTATAATGCTG 415  
Db 301 TTTTGTGATAGATCAGCCGCATATGTTAAATGCAAACTAACTTTTACTATAATGCTG 360  
QY 416 GTTCTCCTTAAGGTGAAGCTCCTGGCGTTTAAGTTTTCGCTGCTGCTGCCGCTGAG 475  
Db 361 GTTCTCCTTAAGGTGAAGCTCCTGGCGTTTAAGTTTTCGCTGCTGCTGCCGCTGAG 420  
QY 476 GTGTTGCTGCCGTTACTAGTTAATGTTGACCTTGCCAACTAAACAAACGATTTCTCCTG 535  
Db 421 GTGTTGCTGCCGTTACTAGTTAATGTTGACCTTGCCAACTAAACAAACGATTTCTCCTG 480  
QY 536 CCACATCGAGTGCCTAAGCTAATTTAGCCACATAATGATAGCANTTAATGTCCTACTGCA 595  
Db 481 CCACATCGAGTGCCTAAGCTAATTTAGCCACATAATGATAGCANTTAATGTCCTACTGCA 540  
QY 596 CTGTACTTGATGATGGAGTGACACTTGTGTTTAAATACATCAGCCACATATATGTTAAAT 655  
Db 541 CTGTACTTGATGATGGAGTGACACTTGTGTTTAAATACATCAGCCACATATATGTTAAAT 600  
QY 656 GCAGACTTAACCTTTTACTATAATGTTGTTCTCCTTAAGTGAAGCTCCTGGCGTTTAA 715  
Db 601 GCAGACTTAACCTTTTACTATAATGTTGTTCTCCTTAAGTGAAGCTCCTGGCGTTTAA 660  
QY 716 TTTTGTGCTGCTGCTGCCGCTGCAGGTGTTGCTGCCGTTACTAGTTAATGTTACCTT 775  
Db 661 TTTTGTGCTGCTGCTGCCGCTGCAGGTGTTGCTGCCGTTACTAGTTAATGTTACCTT 720  
QY 776 GCCAATAAACAACCAACGATTTCTCCTGCACATGAGGTGCCTAAGCTAATTTAGCCACAT 835  
Db 721 GCCAATAAACAACCAACGATTTCTCCTGCACATGAGGTGCCTAAGCTAATTTAGCCACAT 780  
QY 836 AATGAGTACTTAAATGTCACACTGACCTGCAATTCAGACGGAGTGACACTTGTGTTTAA 895  
Db 781 AATGAGTACTTAAATGTCACACTGACCTGCAATTCAGACGGAGTGACACTTGTGTTTAA 840  
QY 896 GTAATTCATCCACATAATGTTCTTAAATGCATTGCTTAATTTACTTTTTTAATGTAATTCG 955  
Db 841 GTAATTCATCCACATAATGTTCTTAAATGCATTGCTTAATTTACTTTTTTAATGTAATTCG 900  
QY 956 AAGCAGGTAAGGTTAATGTTTAAAGTCTCCAGTAAGTAAACTACTCCAGCAGCATGCTC 1015  
Db 901 AAGCAGGTAAGGTTAATGTTTAAAGTCTCCAGTAAGTAAACTACTCCAGCAGCATGCTC 960  
QY 1016 CAGGTAATGCTACTTAAGCCACATAATGTTTGACCACATGCTGCTGGGTACAGTAC 1075  
Db 961 CAGGTAATGCTACTTAAGCCACATAATGTTTGACCACATGCTGCTGGGTACAGTAC 1020



```

Db 1169 GTTGATGATGCTACATCAACTAAATTTGTAGCTTTAGCAAGTGAATGACTACTAAATCTTAG 1228
Qy 1135 GCTGGCTTTTTCATCAAAACAACTGGTTTACAGCAGGTACTGATACATGTAAGTAA 1194
Db 1229 CTAACATTTATGATCAAAACATCTGGTTTGCAGCAGGTACTGATACATGTAAGTAA 1288
Qy 1195 TGTACTAAAAATTAACCTCTGCTGGCCACAGCTAAAGTATATGCTGAAGCTACTCAAAA 1254
Db 1289 TGTCTAAAAATTAACCTCTGCTGGCTGCTAGCTAAAGTATATGCTGAAGCTACTCAAAA 1348
Qy 1255 GTATAATGCGCTCCACTACTTTCGCTAAATTTTATCGATTTCCCTATTATTTATTTCT 1314
Db 1349 GCATTAATGCGCCAGTTCCTACTTTCGCAAAATTTTATCAATGCTTCAATATTTATTTCT 1408
Qy 1315 TTCTATTATTG 1326
Db 1409 TTCTATTATTG 1420

RESULT 4
AF324424
LOCUS
DEFINITION Ichthyophthirius multifiliis 3026 bp DNA linear INV 27-FEB-2002
(IAG52A) gene, complete cds.
ACCESSION AF324424
KEYWORDS
SOURCE
ORGANISM Ichthyophthirius multifiliis.
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Ophryoglenina; Ichthyophthirius.
Lin.Y., Lin.T.L., Wang,C.C., Wang,X., Stieger,K., Klopfleisch,R.
and Clark,T.G.
TITLE Variation in primary sequence and tandem repeat copy number among
i-antigens of Ichthyophthirius multifiliis
Mol. Biochem. Parasitol. 120 (1), 93-106 (2002)
21839613
JOURNAL
MEDLINE
PUBMED 11849709
REFERENCE 2 (bases 1 to 3026)
AUTHORS Lin.Y., Lin.T.-L. and Clark,T.G.
DIRECT SUBMISSION
TITLE Submitted (27-NOV-2000) Microbiology and Immunology, NYSCVM Cornell
University, Ithaca, NY 14853, USA
JOURNAL
FEATURES
source
1..3026
/organism="Ichthyophthirius multifiliis"
/macronuclear
/serotype="D"
/isolate="G5"
/db_xref="taxon:5932"
1285..2817
/gene="IAG52A"
1285..2817
/gene="IAG52A"
/product="immobilization antigen isoform"
1330..2736
/gene="IAG52A"
/note="52 kDa protein"
/codon_start=1
/transl_table=6
/product="immobilization antigen isoform"
/protein_id="AAK01661.1"
/db_xref="GI:12698727"
/translation="MKNNILVILIISLIFNIQKSNCPVGTETNAGVDDILGTANCC
VNCQKFNYYNNAAFVPGASTCTPCPKDQAQNPATANLVQCNVKCPAGTAIA
GGCTDYAAIITTEVCNCRNFYNENAFNAGASTCTACPNRVGALTAGNAATIVAQ
CNVACPTGALDQGVTDYVRSTECVKRLNFIYNGNNGTFFNPGKXSQCTPCPAIK
PANYQATLGNDAITTAQCNVACPDGTISAAGVNNVVAQNETCTNCAPNFYNNAPNF
NPGNSTLPCPNKRDYGAETAGGAATLAKQCNACPDGTAIASGATNVIYLQTECLN
CAANFYDGNFNQGSRRKACPNKRVQAVATAGGTATLIAQCELECPAGVLDGT
TSTYKQAASECVKCAANFYTKQTDWVAGIDICTSCNKKLITSGAENLPESAKNIQC
DFANFLSISLLLSYLL"

```

```

BASE COUNT      1177 a      316 c      355 g      1178 t
ORIGIN
Query Match      19.0%; Score 252.6; DB 3; Length 3026;
Best Local Similarity 56.7%; Pred. No. 4.6e-38;
Matches 660; Conservative 0; Mismatches 394; Indels 111; Gaps 6;
Qy 167 GTGCTGCTTAAGGAGAGCTTAATGGTAATTAACCTTTTCGACGCAAAATATGCTGTAGAG 226
Db 1673 GTGTTAATTCGTAGAAATTAATTTTATATGAATAATGCTCCAAATTTTAATGAGAGGTGCTA 1732
Qy 227 GTATATGCTGATACCATGCCAAATAAAGAGAGTAGGCTCTGTTACCAATGAGAGTACTTAG 286
Db 1733 GTACATGCACAGCTTGTCCGGTAAACAGAGTTGGTGGTGCATTTGCTGCTGTAATGCGG 1792
Qy 287 CTACTTTTAGCCACATAATGCAGTACITTAATGTCCTACTGCACTGCACCTGATGATGAG 346
Db 1793 CTACCATAGTCGCATTAATGTAAGTCGCATGCTCTACTGCTACTGCTACTGATGATGAG 1852
Qy 347 TGACAGATGTTTTGTATAGATGAGCGCATATGTTAAATSCAAAGCTAACTTTTACT 406
Db 1853 TAACTTACTGATTAATGTTAGATCATTCACAGAAATGTTAAATGTAGACTTAACCTTTTACT 1912
Qy 407 ATAATGCTGCTTCTCCCTTAAGGTGAAGCTCCCTGGCGTTTAAAGTTTTCGCTGCTGCTG 466
Db 1913 ATAATGCTGAATGTAATGTAATCTCTTCAATCCAGGTAAGAAAGTTAATGCACACCTTGC 1972
Qy 467 CCCTGCAGGCTGTTGCTGCCGTTACTAGTTAATGTTACCTTGCCAACTAAACAAAAAG 526
Db 1973 CGCAATTAACCTGCT-----TCAATCCAGGTAATAGTACAT 2176
Qy 527 ATTCTCTGCCACTGCAGGTGCTTAAGCTAATTTAGCCACATAATGTAGCAATTAATGTC 586
Db 1994 TTGCTTTAAGCTACTTTAGTGAATGATGCTACAAATAACGCGCATAGTAACTGTCATGCC 2053
Qy 587 CTACTGCACTGTACTTTGATGATGGAGTGACACTTGTTTTAAATACATCAGCCACATTA 646
Db 2054 CTGATGCTACTATAAGTCTGCTGGAGT---AAATAATGGGTAGCACAAAACACTGAAT 2110
Qy 647 GTGTTAAATGCAGACTAACTTTTACTATAATGGTGGTTCCTCTTAAAGGTGAAGCTCCTG 706
Db 2111 GTACTAATTTGCTCTCACTTAACTTTTACAATAATAATGCTCTTAATTT----- 2155
Qy 707 GCGTTTAAAGTTTGTCTGCTGCTGCCGCTGCGAGGTGTTGCTGCCGTACTAGTTAAT 766
Db 2156 -----TCAATCCAGGTAATAGTACAT 2176
Qy 767 GTGTACTTTGCCAAATAAACAACAAAGATTCCTCT---GCCACTGCAGGTGCCCTAAGCTA 823
Db 2177 GCCTACCTTGCCCGACCAATAAAGATTAATGGTGTGAAGCCACTGCAGGTGGTGGCCCTA 2236
Qy 824 ATTTAGCCACATAATGCGAGTACTTAATGTCACATGGCAGCTGCAATTCAGACGAGAGTGA 883
Db 2237 CTTTAGCCAAATAATGTAATATTTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2296
Qy 884 CACTTGTTTTGTAGTAATTCATCCACATAATGTTCTTAATGCAATGCTTAATTTACTTTTA 943
Db 2297 CTAATTAATGTAATTAATAACAGAAATGCTAAATTTGCTGCTGCTGCTGCTGCTGCTG 2356
Qy 944 ATGGTAATTCGAAGCAGGTAAAAGTTAATGTTTAAAGTGTCCAGTAAGTAAAGT---A 1000
Db 2357 GTAATAATTTCTAGCAGGAAGTAGTAGATGCAAAAGCATGTCCAGCAAAATAAAGTTTAAG 2416
Qy 1001 CTCAGCAGCATGCTCCAGGTAACTGCTACTTAAAGCCACATAATGTTGACACATGTC 1060
Db 2417 GCGTGTAGCAACTGCAGGTGGTACTGCTACTTTAATTTGCTGCTGCTGCTGCTGCTGCTG 2476
Qy 1061 CTGCTGGTACAGTACTTGTATGATGGAACATCAACTAAATTTTGTAGCTTCGCGCAACTGA 1120
Db 2477 CTGCTGGTACTGCTACTCACCAGTGGACACATCTACTTTAATAAGCAGCAGTCTGAAT 2536
Qy 1121 GTACTAAATGTTCTGCTGGCTTTTTCATCATCAAAACAACTGGTTTACAGCAGGTACTG 1180

```









Db 102683 CTGCTACTCTCTCCACAGCTACTGCAACTGCTGCTGCTGCT 102640

## RESULT 7

AC099174/c

LOCUS

DEFINITION

AC099174

AC099174.5

GI:21728909

HTG: HTGS\_PHASE1

SOURCE

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 176822)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,

Barbosa,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabrieli,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,

Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,

Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,

Jacobson,B., Jia,X., Johnson,R., Jollivet,S., Joudah,S.,

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,

Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,

Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,

Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,S.,

Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,

Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,

Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojibokan,I., Roife,M., Ruiz,S., Savery,G.,

Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,

Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,

Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,

Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,

Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,

Wu,C., Wu,Y., Wu,F., Zhou,J., Zorrilla,S., Nelson,D.,

Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 176822)

Worley,K.C.

Direct Submission

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 176822)

Worley,K.C.

Direct Submission

Submitted (12-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 11, 2002 this sequence version replaced gi:18846108.

----- Genome Center

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GZV  
Center clone name: CH230-9811  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 115453 bases at least Q40  
Consensus quality: 125501 bases at least Q30  
Consensus quality: 133023 bases at least Q20  
-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 77 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
1052: contig of 1052 bp in length  
1152: gap of unknown length  
1153  
2179: contig of 1027 bp in length  
2180  
2279: gap of unknown length  
2280  
3355: contig of 1076 bp in length  
3356  
3455: gap of unknown length  
3456  
4834: contig of 1379 bp in length  
4835  
4934: gap of unknown length  
4935  
6222: contig of 1288 bp in length  
6223  
6322: gap of unknown length  
6323  
7400: contig of 1078 bp in length  
7401  
7500: gap of unknown length  
7501  
8807: contig of 1307 bp in length  
8808  
8907: gap of unknown length  
8908  
10266: contig of 1359 bp in length  
10267  
10366: gap of unknown length  
10367  
11662: contig of 1296 bp in length  
11663  
11762: gap of unknown length  
11763  
12812: contig of 1050 bp in length  
12813  
14136: contig of 1224 bp in length  
14137  
14236: gap of unknown length  
14237  
15446: contig of 1210 bp in length  
15447  
15546: gap of unknown length  
15547  
16866: contig of 1320 bp in length  
16867  
16966: gap of unknown length  
16967  
18700: contig of 1734 bp in length  
18701  
18800: gap of unknown length  
18801  
20358: contig of 1558 bp in length  
20359  
20458: gap of unknown length  
20459  
22025: contig of 1567 bp in length  
22026  
22125: gap of unknown length  
22126  
24515: contig of 2390 bp in length  
24516  
25846: gap of unknown length  
25847  
25946: contig of 1231 bp in length  
25947  
27176: contig of 1230 bp in length  
27177  
27276: gap of unknown length  
27277  
28283: contig of 1007 bp in length  
28284  
29995: contig of 1612 bp in length  
29996  
30095: gap of unknown length  
30096  
31459: contig of 1364 bp in length  
31460  
31559: gap of unknown length  
31560  
32924: contig of 1365 bp in length  
32925  
34073: gap of unknown length  
34074  
34173: gap of unknown length

[illegible]

```

Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175235 bases at least Q40
Consensus quality: 176522 bases at least Q30
Consensus quality: 177171 bases at least Q20
Insert size: 198000; agarose-fp
Insert size: 178453; sum-of-contigs
Quality coverage: 4.48 in Q20 bases; agarose-fp
Quality coverage: 4.99 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	7554:	contig of 7554 bp in length
*	7555	gap of unknown length
*	7635	contig of 8695 bp in length
*	16350	gap of unknown length
*	16350	contig of 10026 bp in length
*	26475:	gap of unknown length
*	26476	contig of 9305 bp in length
*	26576	gap of unknown length
*	35881	contig of 26030 bp in length
*	35981	gap of unknown length
*	62011	contig of 28455 bp in length
*	62111	gap of unknown length
*	90565:	contig of 26518 bp in length
*	90566	gap of unknown length
*	117183:	contig of unknown length
*	117184	gap of unknown length
*	117284	contig of 45736 bp in length
*	163019:	gap of unknown length
*	163020	contig of 3819 bp in length
*	163120	gap of unknown length
*	166939	contig of 2274 bp in length
*	167039	gap of unknown length
*	169313	contig of 2227 bp in length
*	169413	gap of unknown length
*	171640	contig of 7814 bp in length.
*	171740	gap of unknown length

```

/note="assembly_name:Contig9"
BASE COUNT      48665 a 37751 c 36838 g 54991 t 1108 others
ORIGIN

Query Match              7.0%: Score 93: DB 2: Length 179553;
Best Local Similarity 44.1%: Pred. No. 3e-08;
Matches 537: Conservative 0: Mismatches 667: Indels 15: Gaps

QY  30  TATTTCCTTTATTTATTAATGAATTAAGAGCTGTTCCTATGCTCTGACTGTAATAGACTCA 89
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174755 TATCGTGTCTTTTGTATTGTGTCTGTGTGTGTGCTGCTGCTACTACTGTGTGTGTGCTGC 174814
QY  90  AGCTGGATTGACTGATGATGAGTGTGCTGATCTTGGTACTTGTGTAAATTCGACACCTAA 149
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174815 TGTGCTCTTATGTGTGCTGCTGCTGTGTGTGTGTGTGCTGCTGCTATCTTGTCTCTGT 174874
QY  150  TTTTCTACTATAATGTGTGCTGCTTAAAGGAGAAGCTAAATGGTAATTAACCTTTTCGCAGC 209
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174875 TGTGTCTGATGCTGCTACTGTTGCTATTGCTGTGTGCTGCTACTATTACTGTTCTATTGC 174934
QY  210  AAATAATGCTCTAGAGGTATATGTATACCATGCCAATAAACAGAGTAGGCTGTGTTCAC 269
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174935 TCGTCTACTGCTGCTGTGTGTGCTGCTACTGTTGTGCTACTATGGTAATATTGCTGTC 174994
QY  270  CAATGACGGTGACTTTAGCTACTTTTAGCCACATAATGACGACTTAATGTCTACTCTGCAC 329
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 174995 TGTCTGCTGCTACTGTTGTACTGTGTAATTCGTTGCTACTGCTGCTGCTGCTGTTGTTAA 175054
QY  330  TGCACCTGATGATGAGTGACAGATGTTTTGTATAGATCAGCCGCATAAATGTGTTAAATG 389
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175055 TTGTGATGCTATTGCTGTGTGCTGATGCTTTTGTGCTATTGCTGTGCTGCTGCTGATAC 175114
QY  390  CAAACCTAACTTTTACTATATAATGTGTGCTCTCTTAGGTTGAAGCTCCCTGGCTTTTAACT 449
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175115 TCTTGTATTGCTTTTGTGTTGCTGCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 175174
QY  450  TTTTGCTGCTGGTGTGCGCGCTGCAGGTGTGCTGCGGTTTACTAGTAAATGTGTACTCTTG 509
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175175 TGTGTGCTGCTCTTATTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175234
QY  510  CCAACTAAACAAAACGATTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 569
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175235 TACTGCTGCTATTGTTGCTGTTGCTGATGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTG 175291
QY  570  ATGTAGCAATTAATGTCCTACTGSCACTGTACTGTATGATGGAGTGACACTGTTGTTTTAA 629
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175292 ---TATCGTGTCTTTGTTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175348
QY  630  TACATCAGCCACATTATGCTTAAATGCCAGACCTTAATTTTACTATATATGTTGTTGCTGCC 689
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175349 TGTGCTGCTCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175408
QY  690  TTAAGGTGAAGCTCCTGGCGCTTTAAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 749
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175409 TATTGCTGTTGCTGCTACTATTACTGTGCTATTGCTGCTACTATTGCTGCTGCTGCTGCTG 175468
QY  750  TGGCGTTTACTAGTTAATGTACTCTTGGCCAAATAAACAAAAACGATTCCTGTCACACTGC 809
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175469 TG-----CTACTGTTGTGCTACTATGTAATGATTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175522
QY  810  AGGTGCTTAAGCTAAATTAAGCCACATAATGCACTACTTAATGTCACACTGGCAGCTGCAAT 869
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175523 TATTACTGCGATTTCTGTTGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175582
QY  870  TCAAGCGGAGTGACACTGTTTTTGTAAATTCATCCACATAAATGTTCTTAATGCAATTCG 929
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175583 TGTTAATTCGTGATACTATTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 175642
QY  930  TAATTACTTTTTTAATGGTAATTTTCAAGCAGGTAAAGTTAATGTTTAAAGTGTCCAGT 989
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175643 TGTGNCATATTGCGCTTTTGTAAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 175702
QY  990  AAGTAAACACTCTCCAGCACATGCTCCAGGTAAATCTGCTACTTAAAGCCACATAATGTTT 1049
    |||  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```







Db 8155 TTGTTGCTGCTTTATTGTTGCTGCTGCTGTTGCTGTTGTTGCTGCTGCTGTTGCTGTT 8096  
Qy 509 CCCAACTAAACAAAAGATTCCTCCGCACTGCAGTGCCTAAGCTAATATTTAGCCACAT 568  
Db 8095 TTACTGCTGCTATTTGCTGTTGTTGCTGATGCTGTTGCTGTTGCTGATGCTGTTGCTG 8038  
Qy 569 AATGTAGCAATTAATGTCCTACTGGCAGCTGTACTTGTATGATGGAGTGACACTTGTGTTTA 628  
Db 8037 ----TATCGTTGCTTTTATTGTTGCTGTTGCTGTTGCTGCTGCTGCTGCTGTTGTTG 7982  
Qy 629 ATACATAGCCACATATATGTTTAATGAGCAGACCTAACTTTTACATATAATGTTGTTCTC 688  
Db 7981 CTGCTGCTCTCTATTTGTTGCTGCTGCTGTTGCTGTTGCTGTTGCTGATGCTGCTACTGTTG 7922  
Qy 689 CTTAAGGTGAAGCTCCTGGGCTTTAAGTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 808  
Db 7921 CTATTGCTGTTGCTACTATTTACTGTTGCTATTTGCTGCTGCTGCTGCTGCTGCTGCTG 748  
Qy 749 CTGCGGTTACTAGTTAATGTGTACCTTGCCAAATAAACAACAAAGATTCTCTGCTGCTG 808  
Db 7861 CTG-----CTACTGTTGTTGCTACTATGTTAATGATTGCTGCTGCTGCTGCTGCTGCTG 7808  
Qy 809 CAGGTGCTAAGCTAATTTAGCCACATATGCACTACTTAATGTCCAACTGCACTGCA 868  
Db 7807 TTATTACTGGATTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7748  
Qy 869 TTCAAGACGAGTGACACTGTTTTCGTAATTCATCCACATAAATGTTCTTAATGCTATG 928  
Db 7747 TTGTTAATTCGATACTATTGCTGTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTA 7688  
Qy 929 CTAATTACTTTTTTAATGTAATTCGAAGCAGGTAAGTTAATGTTAAGTGCTCCAG 988  
Db 7687 CTGTTGCTATGTTGCTGTTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7628  
Qy 989 TAAGTAAACTACTCCAGCACATGTCAGGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048  
Db 7627 TGCTGCTATTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7571  
Qy 1049 TGACCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1108  
Db 7570 TTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7511  
Qy 1109 CGCAACTCAATGCTACTAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1168  
Db 7510 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7451  
Qy 1169 CAGCAGGTACTGATCATGCTACTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1228  
Db 7450 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7391  
Qy 1229 AAGTATGCTGGAAGCTACT 1248  
Db 7390 CTGTTATTGCTGCGACTTCT 7371

RESULT 11  
AC119627  
LOCUS Rattus norvegicus clone CH230-62L12, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 70 unordered pieces.  
ACCESSION AC119627  
VERSION AC119627.2 GI:21746574  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE 1 (bases 1 to 189461)  
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homands,J., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okunodu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,  
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,  
Wu,C., Wu,Y., Zhou,J., Zorrilla,S., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 189461)  
Worley,K.C.  
Direct Submission  
Submitted (30-APR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 189461)  
Worley,K.C.  
Direct Submission  
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20340348.  
-----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
-----  
Center project name: GUQS  
Center clone name: CH230-62L12  
-----  
Summary Statistics  
Sequencing vector: Plasmid:  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 134624 bases at least Q40  
Consensus quality: 141279 bases at least Q30  
Consensus quality: 145103 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 70 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as









AUTHORS  
TITLE  
JOURNAL  
  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

McPherson,J.D. and Waterston,R.H.  
Direct Submission  
Submitted (15-FEB-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 204259)  
McPherson,J.D. and Waterston,R.H.  
Direct Submission  
Submitted (29-JUL-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
On May 23, 2002 this sequence version replaced gi:18874253.

COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@watson.wustl.edu  
----- Project Information -----  
Center project name: M\_BA0155B20  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Sequencing method: plasmid; 100%  
Chemistry: Dye-primer; 0% of reads  
Chemistry: Dye-terminator; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 198852 bases at least Q40  
Consensus quality: 200032 bases at least Q30  
Consensus quality: 201011 bases at least Q20  
Insert size: 211000; agarose-fp  
Insert size: 203242; sum-of-contigs  
Quality coverage: 8.38 in Q20 bases; agarose-fp  
Quality coverage: 8.78 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 16 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1226: contig of 1226 bp in length  
\* 1227 1326: gap of unknown length  
\* 1327 2375: contig of 1049 bp in length  
\* 2376 2475: gap of unknown length  
\* 2476 3800: contig of 1325 bp in length  
\* 3801 5876: contig of 1976 bp in length  
\* 5877 5977: gap of unknown length  
\* 5977 7870: contig of 1893 bp in length  
\* 7870 7969: gap of unknown length  
\* 7970 9641: contig of 1672 bp in length  
\* 9642 14703: contig of 4962 bp in length  
\* 14704 14804: gap of unknown length  
\* 14804 20620: contig of 5817 bp in length  
\* 20621 20720: gap of unknown length  
\* 20721 30389: contig of 9669 bp in length  
\* 30390 30489: gap of unknown length  
\* 30490 38838: contig of 8349 bp in length  
\* 38839 38939: gap of unknown length  
\* 38939 51520: contig of 12581 bp in length  
\* 51520 51620: gap of unknown length  
\* 51620 70099: contig of 18479 bp in length  
\* 70099 70199: gap of unknown length  
\* 70199 100939: contig of 30741 bp in length  
\* 100940 101039: gap of unknown length  
\* 101040 131732: contig of 30693 bp in length  
\* 131733 131832: gap of unknown length  
\* 131833 163882: contig of 32049 bp in length  
\* 163882 163982: gap of unknown length  
\* 163982 204259: contig of 40278 bp in length.

Location/Qualifiers

source

1..204259  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP23-155B20"  
1..1226  
/note="assembly\_name:Contig13"  
1327..2375  
/note="assembly\_name:Contig15"  
2476..3800  
/note="assembly\_name:Contig16"  
3901..5876  
/note="assembly\_name:Contig17"  
5977..7869  
/note="assembly\_name:Contig18"  
7970..9641  
/note="assembly\_name:Contig19"  
9742..14703  
/note="assembly\_name:Contig20"  
14804..20620  
/note="assembly\_name:Contig21"  
20721..30389  
/note="assembly\_name:Contig22"  
30490..38838  
/note="assembly\_name:Contig23"  
38939..51519  
/note="assembly\_name:Contig24"  
51620..70098  
/note="assembly\_name:Contig25"  
70199..100939  
/note="assembly\_name:Contig26"  
101040..131732  
/note="assembly\_name:Contig27"  
131833..163881  
/note="assembly\_name:Contig28"  
163982..204259  
/note="assembly\_name:Contig29"  
BASE COUNT 61509 a 41663 c 41736 g 57836 t 1515 others  
ORIGIN

Query Match 6.6%; Score 87.6; DB 2; Length 204259;  
Best Local Similarity 44.0%; Pred. No. 3e-07;  
Matches 369; Conservative 0; Mismatches 469; Indels 0; Gaps 0;  
Qy 411 TGGTGGTCTCTTAAGTGAAGCTCCTGGCGTTTAAAGTTTGGTCTGCTGGTGGCCG 470  
Db 179959 TGCTTCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180018  
Qy 471 TGCAGGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530  
Db 180019 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180078  
Qy 531 TCTTGCCACTGCAAGTGCCTTAAGCTAATTTAGCCACATAATGTAGCAATTAATGCTCTAC 590  
Db 180079 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180138  
Qy 591 TGGCACTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650  
Db 180139 TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180198  
Qy 651 TAAATGCAGACCTAACTTTTACTATATATATATATATATATATATATATATATATAT 710  
Db 180199 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180258  
Qy 711 TTAAGTTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 770  
Db 180259 TGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180318  
Qy 771 ACCTTGCCAAATAAACAACCAATTTCTCTGCCACTGCAAGTGGCTTAAGCTAATTAAG 830  
Db 180319 TTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180378  
Qy 831 CACATAATGCAGTACTTAATATGTCCAACTGCAACTGCAATTTCAAGACGGAGTACACTTGT 890





